Boris Tsirelson
Tel Aviv University

Scaling Limit, Noise, Stability

Summer school for probability
Saint Flour, France
July 2002

Contents

1 A first look

1a Foreword

Functions of many independent random variables are a tenor of probability theory. Some examples follow.

- Classical limit theorems investigate linear functions, such as

$$f(\xi_1, \dots, \xi_n) = \frac{\xi_1 + \dots + \xi_n}{\sqrt{n}}$$

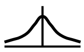
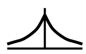
  for $n \to \infty$. Functional limit theorems lead to Brownian motions.
- Stochastic differential equations represent a given diffusion process as a function of a Brownian motion, or a white noise. The latter may be treated as an infinitely divisible reservoir of independent random variables.
- Percolation theory investigates some very special nonlinear functions of independent two-valued random variables, either in the limit of an infinite discrete lattice, or in the scaling limit (especially successful nowadays).

Though, percolation is of dual citizenship: probability, and statistical physics. The following example belongs rather to computer science:

- Stability and sensitivity of Boolean functions, especially, properties of large random graphs.

We'll see later that stability and sensitivity are quite important for probability theory.

In many cases we have a sequence of "more elementary" probabilistic models, and we want to construct a "less elementary" model by a limiting procedure. Some examples follow.

- Brownian motion is a scaling limit of random walks.
- Striving to understand turbulence, one may start with a stochastic flow whose correlation function is smoothed out, and look what happens when the smoothing disappears gradually.  $\to$ 
- Percolation theory strives to a conformally invariant scaling limit for discrete models of percolation.

The most interesting thing is a scaling limit as a transition from a lattice model to a continuous model. A transition from a finite sequence to an infinite sequence is much simpler, but still nontrivial, as we'll see on simple toy models.

1b Two toy models

Classical theorems about independent increments are exhaustive, but a small twist may surprise us. I'll demonstrate the twist on two models, 'discrete' and 'continuous'. The 'continuous' model is a Brownian motion on the circle. The 'discrete' model takes on two values ± 1 only, and increments are treated multiplicatively: $X(t)/X(s)$ instead of the usual $X(t) - X(s)$. Or equivalently, the 'discrete' process takes on its values in the two-element group $\mathbb{Z}_2$; using additive notation we have $\mathbb{Z}_2 = \{0, 1\}$, $1 + 1 = 0$, increments being $X(t) -$

$X(s)$. In any case, the twist stipulates values in a compact group (the circle, $\mathbb{Z}_2$, etc.), in contrast to the classical theory, where values are in $\mathbb{R}$ (or another linear space). Also, the classical theory assumes continuity (in probability), while our twist does not. The 'continuous' process (in spite of its name) is discontinuous at a single instant $t = 0$. The 'discrete' process is discontinuous at $t = \frac{1}{n}$, $n = 1, 2, \dots$, and also at $t = 0$; it is constant on $[\frac{1}{n+1}, \frac{1}{n})$ for every n.

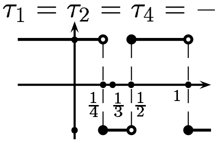
1b1 Example. Introduce an infinite sequence of random signs $\tau_1, \tau_2, \dots$; that is,

$$\mathbb{P}\big(\tau_k = -1\big) = \mathbb{P}\big(\tau_k = +1\big) = \frac{1}{2} \quad \text{for each } k,$$
$$\tau_1, \tau_2, \dots \quad \text{are independent.}$$

For each n we define a stochastic process $X_n(\cdot)$, driven by $\tau_1, \dots, \tau_n$, as follows:

$$X_n(t) = \prod_{k:\frac{1}{n} \le \frac{1}{k} \le t} \tau_k \, .$$

a sample path of X_4
(here $\tau_1 = \tau_2 = \tau_4 = -1, \tau_3 = +1$)



For $n \to \infty$, finite-dimensional distributions of X_n converge to these of a process $X(\cdot)$. Namely, X consists of countably many random signs, situated on intervals $[\frac{1}{k+1}, \frac{1}{k})$. Almost surely, X has no limit at $0+$. We have

$$\frac{X(t)}{X(s)} = \prod_{k:s<\frac{1}{k}\le t} \tau_k \tag{1b2}$$

whenever $0 < s < t < \infty$. However, it does not hold when $s < 0 < t$. Here, the product contains infinitely many factors and diverges almost surely; nevertheless, the increment $X(t)/X(s)$ is well-defined. Each X_n satisfies (1b2) for all s, t (including $s < 0 < t$; of course, $k \le n$), but X does not. Still, X is an independent increment process (multiplicatively); that is, $X(t_2)/X(t_1), \dots, X(t_n)/X(t_{n-1})$ are independent whenever $-\infty < t_1 < \dots < t_n < \infty$. However, we cannot describe the whole X by a countable collection of its independent increments. The infinite sequence of $\tau_k = X(\frac{1}{k}+)/X(\frac{1}{k}-)$ does not suffice since, say, $X(1)$ is independent of $(\tau_1, \tau_2, \dots)$. Indeed, the global sign change $x(\cdot) \mapsto -x(\cdot)$ is a measure-preserving transformation that leaves all τ_k invariant. The conditional distribution of $X(\cdot)$ given $\tau_1, \tau_2, \dots$ is concentrated at two functions of opposite global sign. It may seem that we should add to $(\tau_1, \tau_2, \dots)$ one more random sign τ_∞ independent of $(\tau_1, \tau_2, \dots)$ such that $X(\frac{1}{k})$ is a measurable function of $\tau_k, \tau_{k+1}, \dots$ and τ_∞. However, it is impossible. Indeed, $X(1) = \tau_1 \dots \tau_k X(\frac{1}{k})$. Assuming $X(\frac{1}{k}) = f_k(\tau_k, \tau_{k+1}, \dots; \tau_\infty)$ we get $f_1(\tau_1, \tau_2, \dots; \tau_\infty) = \tau_1 \dots \tau_{k-1} f_k(\tau_k, \tau_{k+1}, \dots; \tau_\infty)$ for all k. It follows that $f_1(\tau_1, \tau_2, \dots; \tau_\infty)$ is orthogonal to all functions of the form $g(\tau_1, \dots, \tau_n)h(\tau_\infty)$ for all n, thus, to a dense (in L_2) set among all functions of $\tau_1, \tau_2, \dots; \tau_\infty$; a contradiction.

So, for each n the process X_n is driven by (τ_k), but the limiting process X is not.

1b3 Example. (See also [1].) We turn to the other, 'continuous' model. For any $\varepsilon \in (0,1)$ we introduce a (complex-valued) stochastic process

$$Y_\varepsilon(t) = \begin{cases} \exp\bigl(iB(\ln t) - iB(\ln \varepsilon)\bigr) & \text{for } t \ge \varepsilon, \\ 1 & \text{otherwise,} \end{cases}$$

where $B(\cdot)$ is the usual Brownian motion; or rather, $\bigl(B(t)\bigr)_{t\in[0,\infty)}$ and $\bigl(B(-t)\bigr)_{t\in[0,\infty)}$ are two independent copies of the usual Brownian motion. Multiplicative increments $Y_\varepsilon(t_2)/Y_\varepsilon(t_1), \dots, Y_\varepsilon(t_n)/Y_\varepsilon(t_{n-1})$ are independent whenever $-\infty < t_1 < \dots < t_n < \infty$, and the distribution of $Y_\varepsilon(t)/Y_\varepsilon(s)$ does not depend on ε as far as $\varepsilon < s < t$ (in fact, the distribution depends on t/s only). The distribution of $Y_\varepsilon(1)$ converges for $\varepsilon \to 0$ to the uniform distribution on the circle $|z| = 1$. The same for each $Y_\varepsilon(t)$. It follows easily that, when $\varepsilon \to 0$, finite dimensional distributions of Y_ε converge to these of some process Y. For every $t > 0$, $Y(t)$ is distributed uniformly on the circle; and Y is an independent increment process (multiplicatively); and $Y(t) = 1$ for $t \le 0$. Almost surely, $Y(\cdot)$ is continuous on $(0,\infty)$, but has no limit at $0+$. We may define $B(\cdot)$ by

$$\begin{gathered} Y(t) = Y(1)\exp\bigl(iB(\ln t)\bigr) \quad \text{for } t \in \mathbb{R}, \\ B(\cdot) \quad \text{is continuous on } \mathbb{R}, \end{gathered}$$

then B is the usual Brownian motion, and

$$\frac{Y(t)}{Y(s)} = \frac{\exp(iB(\ln t))}{\exp(iB(\ln s))} \quad \text{for } 0 < s < t < \infty\,.$$

However, $Y(1)$ is independent of $B(\cdot)$. Indeed, the global phase change $y(\cdot) \mapsto e^{i\alpha}y(\cdot)$ is a measure preserving transformation that leaves $B(\cdot)$ invariant. The conditional distribution of $Y(\cdot)$ given $B(\cdot)$ is concentrated on a continuum of functions that differ by global phase (distributed uniformly on the circle). Similarly to the 'discrete' example, we cannot introduce a random variable $B(-\infty)$ independent of $B(\cdot)$, such that $Y(t)$ is a function of $B(-\infty)$ and increments of $B(r)$ for $-\infty < r < \ln t$.

So, for each ε, the process Y_ε is driven by the Brownian motion, but the limiting process Y is not.

1c Our limiting procedures

Imagine a sequence of elementary probabilistic models such that the n-th model is driven by a finite sequence $(\tau_1, \dots, \tau_n)$ of random signs (independent, as before). A limiting procedure may lead to a model driven by an infinite sequence $(\tau_1, \tau_2, \dots)$ of random signs. However, it may also lead to something else, as shown in 1b. That is an occasion to ask ourselves: what do we mean by a limiting procedure?

The n-th model is naturally described by the finite probability space $\Omega_n = \{-1,+1\}^n$ with the uniform measure. A prerequisite to any limiting procedure is some structure able to join these Ω_n somehow. It may be a sequence of 'observables', that is, functions on the disjoint union,

$$f_k : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to \mathbb{R}\,.$$

1c1 Example. Let $f_k(\tau_1,\dots,\tau_n) = \tau_k$ for $n \ge k$. Though, f_k is defined only on $\Omega_k \uplus \Omega_{k+1} \uplus \dots$, but it is enough. For every k, the joint distribution of $f_1,\dots,f_k$ on Ω_n has a limit for $n \to \infty$ (moreover, the distribution does not depend on n, as far as $n \ge k$). The limiting procedure should extend each f_k to a new probability space Ω such that the joint distribution of $f_1,\dots,f_k$ on Ω_n converges for $n \to \infty$ to their joint distribution on Ω. Clearly, we may take the space of infinite sequences $\Omega = \{-1,+1\}^\infty$ with the product measure, and let f_k be the k-th coordinate function.

1c2 Example. Still $f_k(\tau_1,\dots,\tau_n) = \tau_k$ (for $n \ge k \ge 1$), but in addition, the product $f_0(\tau_1,\dots,\tau_n) = \tau_1 \dots \tau_n$ is included. For every k, the joint distribution of $f_0, f_1, \dots, f_k$ on Ω_n has a limit for $n \to \infty$; in fact, the distribution does not depend on n, as far as $n > k$ (this time, not just $n \ge k$). Thus, in the limit, $f_0, f_1, f_2, \dots$ become independent random signs. The functional dependence $f_0 = f_1 f_2 \dots$ holds for each n, but disappears in the limit. We still may take $\Omega = \{-1,+1\}^\infty$, however, f_0 becomes a new coordinate.

That is instructive; the limiting model depends on the class of 'observables'.

1c3 Example. Let $f_k(\tau_1,\dots,\tau_n) = \tau_k \dots \tau_n$ for $n \ge k \ge 1$. In the limit, f_k become independent random signs. We may define τ_k in the limiting model by $\tau_k = f_k / f_{k+1}$; however, we cannot express f_k in terms of τ_k. Clearly, it is the same as the 'discrete' toy model of 1b.

The second and third examples are isomorphic. Indeed, renaming f_k of the third example to g_k (and retaining f_k of the second example) we have

$$g_k = \frac{f_0}{f_1 \dots f_{k-1}}\,; \qquad f_k = \frac{g_k}{g_{k+1}} \text{ for } k > 0\,, \quad \text{and} \quad f_0 = g_1\,;$$

these relations hold for every n (provided that the same $\Omega_n = \{-1,+1\}^n$ is used for both examples); naturally, they give us an isomorphism between the two limiting models.

That is also instructive; some changes of the class of 'observables' are essential, some are not.

It means that the sequence (f_k) is not really the structure responsible for the limiting procedure. Rather, f_k are generators of the relevant structure. The second and third examples differ only by the choice of generators for the same structure. In contrast, the first example uses a different structure. So, what is the mysterious structure?

I can describe the structure in two equivalent ways. Here is the first description. In the commutative Banach algebra $l_\infty(\Omega_1 \uplus \Omega_2 \uplus \dots)$ of all bounded functions on the disjoint union, we select a subset C (its elements will be called observables) such that

(1c4) $\quad C$ is a separable closed subalgebra of $l_\infty(\Omega_1 \uplus \Omega_2 \uplus \dots)$ containing the unit.

In other words,

$$\begin{gathered} C \text{ contains a sequence dense in the uniform topology;} \\ f_n \in C,\ f_n \to f \text{ uniformly} \quad \Longrightarrow \quad f \in C\,; \\ f, g \in C,\ a, b \in \mathbb{R} \quad \Longrightarrow \quad af + bg \in C\,; \\ \mathbf{1} \in C\,; \\ f, g \in C \quad \Longrightarrow \quad fg \in C \end{gathered} \tag{1c5}$$

(here $\mathbf{1}$ stands for the unity, $\mathbf{1}(\omega) = 1$ for all ω). Or equivalently,

(1c6)
$$\begin{gathered} C \text{ contains a sequence dense in the uniform topology;} \\ f_n \in C,\, f_n \to f \text{ uniformly} \quad \Longrightarrow \quad f \in C\,; \\ f, g \in C,\, \varphi : \mathbb{R}^2 \to \mathbb{R} \text{ continuous} \quad \Longrightarrow \quad \varphi(f,g) \in C\,. \end{gathered}$$

You see, on one hand, both $af+bg$ and fg (and $\mathbf{1}$) are special cases of $\varphi(f,g)$. On the other hand, every continuous function on a bounded subset of $\mathbb{R}^2$ can be uniformly approximated by polynomials. The same for $\varphi(f_1,\dots,f_n)$ where $f_1,\dots,f_n \in C$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$ a continuous function. Another equivalent set of conditions is also well-known:

(1c7)
$$\begin{gathered} C \text{ contains a sequence dense in the uniform topology;} \\ f_n \in C,\, f_n \to f \text{ uniformly} \quad \Longrightarrow \quad f \in C\,; \\ f, g \in C,\, a, b \in \mathbb{R} \quad \Longrightarrow \quad af + bg \in C\,; \\ \mathbf{1} \in C\,; \\ f \in C \quad \Longrightarrow \quad |f| \in C\,; \end{gathered}$$

here $|f|$ is the pointwise absolute value, $|f|(\omega) = |f(\omega)|$.

The smallest set C satisfying these (equivalent) conditions (1c4)–(1c7) and containing all given functions f_k is, by definition, generated by these f_k.

Recall that C consists of functions defined on the disjoint union of finite probability spaces Ω_n; a probability measure P_n is given on each Ω_n. The following condition is relevant:

(1c8)
$$\lim_{n\to\infty} \int_{\Omega_n} f\, dP_n \text{ exists for every } f \in C\,.$$

Assume that C is generated by given functions f_k. Then the property (1c8) of C is equivalent to such a property of functions f_k:

(1c9)
For each k, the joint distribution of $f_1,\dots,f_k$ on Ω_n weakly converges, when $n \to \infty$.

Indeed, (1c9) means convergence of $\int \varphi(f_1,\dots,f_k)\, dP_n$ for every continuous function $\varphi : \mathbb{R}^k \to \mathbb{R}$. However, functions of the form $f = \varphi(f_1,\dots,f_k)$ (for all k, φ) belong to C and are dense in C.

We see that (1c9) does not depend on the choice of generators f_k of a given C.

The second (equivalent) description of our structure is the 'joint compactification' of $\Omega_1, \Omega_2, \dots$ I mean a pair (K, α) such that

(1c10)
$$\begin{gathered} K \text{ is a metrizable compact topological space,} \\ \alpha : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to K \text{ is a map,} \\ \text{the image } \alpha(\Omega_1 \uplus \Omega_2 \uplus \dots) \text{ is dense in } K. \end{gathered}$$

Every joint compactification (K, α) determines a set C satisfying (1c4). Namely,

$$C = \alpha^{-1}\big(C(K)\big)\,;$$

that is, observables $f \in C$ are, by definition, functions of the form

$$f = g \circ \alpha, \text{ that is, } f(\omega) = g(\alpha(\omega)), \quad g \in C(K).$$

You see, the Banach algebra C is the same as the Banach algebra $C(K)$ of all continuous functions on K.

Every C satisfying (1c4) corresponds to some joint compactification. Indeed, C is generated by some f_k such that $|f_k(\omega)| \le 1$ for all k, ω. We introduce

$$\alpha(\omega) = \big(f_1(\omega), f_2(\omega), \dots\big) \in [-1,1]^\infty,$$
$$K \text{ is the closure of } \alpha(\Omega_1 \uplus \Omega_2 \uplus \dots) \text{ in } [-1,1]^\infty;$$

clearly, (K, α) is a joint compactification. Coordinate functions on K generate $C(K)$, therefore f_k generate $\alpha^{-1}\big(C(K)\big)$, hence $\alpha^{-1}\big(C(K)\big) = C$.

Finiteness of each Ω_n is not essential. The same holds for arbitrary probability spaces $(\Omega_n, \mathcal{F}_n, P_n)$. Of course, instead of $l_\infty(\Omega_1 \uplus \Omega_2 \uplus \dots)$ we use $L_\infty(\Omega_1 \uplus \Omega_2 \uplus \dots)$, and the map $\alpha : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to K$ must be measurable. It sends the given measure P_n on Ω_n into a measure $\alpha(P_n)$ (denoted also by $P_n \circ \alpha^{-1}$) on K. If measures $\alpha(P_n)$ weakly converge, we get the limiting model (Ω, P) by taking $\Omega = K$ and $P = \lim_{n\to\infty} \alpha(P_n)$.

1d Examples of high symmetry

1d1 Example. Let Ω_n be the set of all permutations $\omega : \{1, \dots, n\} \to \{1, \dots, n\}$, each permutation having the same probability $(1/n!)$;

$$f : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to \mathbb{R} \text{ is defined by}$$
$$f(\omega) = \big|\{k : \omega(k) = k\}\big|;$$

that is, the number of fixed points of a random permutation. Though, f is not bounded, which happens quite often; in order to embed it into the framework of 1c, we make it bounded by some homeomorphism from $\mathbb{R}$ to a bounded interval (say, $\omega \mapsto \arctan f(\omega)$). The distribution of $f(\cdot)$ on Ω_n converges (for $n \to \infty$) to the Poisson distribution $P(1)$. Thus, the limiting model exists; however, it is scanty; just $P(1)$.

We may enrich the model by introducing

$$f_u(\omega) = \big|\{k < un : \omega(k) = k\}\big|;$$

for instance, $f_{0.5}(\cdot)$ is the number of fixed points among the first half of $\{1, \dots, n\}$. The parameter u could run over $[0,1]$, but we need a countable set of functions; thus we restrict u to, say, rational points of $[0,1]$. Now the limiting model is the Poisson process.

Each finite model here is invariant under permutations. Functions f_u seem to break the invariance, but it survives in their increments, and turns in the limit into invariance of the Poisson process (or rather, its derivative, the point process) under all measure preserving transformations of $[0,1]$.

Note also that *independent* increments in the limit emerge from *dependent* increments in finite models.

We feel that all these $f_u(\cdot)$ catch only a small part of the information contained in the permutation. You may think about more information, say, cycles of length $1, 2, \dots$ (and what about length $n/2$?)

1d2 Example. Let Ω_n be the set of all graphs over $\{1, \dots, n\}$. That is, each $\omega \in \Omega_n$ is a subset of the set $\binom{\{1,\dots,n\}}{2}$ of all unordered pairs (treated as edges, while $1, \dots, n$ are vertices); the probability of ω is $p_n^{|\omega|}(1-p_n)^{n(n-1)/2-|\omega|}$, where $|\omega|$ is the number of edges. That is, every edge is present with probability p_n, independently of others. Define $f(\omega)$ as the number of isolated vertices. The limiting model exists if (and only if) there exists a limit $\lim_n n(1-p_n)^{n-1} = \lambda \in [0, \infty)$;[1] the Poisson distribution $P(\lambda)$ exhausts the limiting model.

A Poisson process may be obtained in the same way as before.

You may also count small connected components more complicated than single points.

Note that the finite model contains a lot of independence (namely, $n(n-1)/2$ independent random variables); the limiting model (Poisson process) also contains a lot of independence (namely, independent increments); however, we feel that independence is not inherited here. Rather, the independence of finite models is lost in the limiting procedure, and a new independence emerges.

1d3 Example. Let $\Omega_n = \{-1, +1\}^n$ with the uniform measure, and $f_n : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to \mathbb{R}$ is defined by

$$f_u(\omega) = \frac{1}{\sqrt{n}} \sum_{k<un} \tau_k(\omega) ;$$

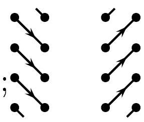
as before, $\tau_1, \dots, \tau_n$ are the coordinates, that is, $\omega = \big(\tau_1(\omega), \dots, \tau_n(\omega)\big)$; and u runs over rational points of $[0, 1]$. The limiting model is the Brownian motion, of course.

Similarly to Example 1d1, each finite model is invariant under permutations. The invariance survives in increments of functions f_k, and in the limit, the white noise (the derivative of the Brownian motion) is invariant under all measure preserving transformations of $[0, 1]$.

1e Example of low symmetry

Example 1d3 may be rewritten via the composition of random maps

$$\alpha_-, \alpha_+ : \mathbb{Z} \to \mathbb{Z} ,$$
$$\alpha_-(k) = k - 1 , \quad \alpha_+(k) = k + 1 ;$$
$$\alpha_\omega = \alpha_{\tau_n(\omega)} \circ \dots \alpha_{\tau_1(\omega)} ;$$



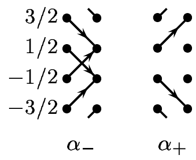
α_- α_+

thus, $\alpha_\omega(k) = k + \tau_1(\omega) + \dots + \tau_n(\omega)$, and we may define $f_1(\omega) = \frac{1}{\sqrt{n}} \alpha_\omega(0)$, which conforms to 1d3. Similarly, $f_u(\omega) = \frac{1}{\sqrt{n}} \alpha_{\omega,u}(0)$, where $\alpha_{\omega,u}$ is the composition of $\alpha_{\tau_k(\omega)}$ for $k \le un$. The order does not matter, since α_-, α_+ commute, that is, $\alpha_- \circ \alpha_+ = \alpha_+ \circ \alpha_-$. It is interesting to try a pair of noncommuting maps.

[1] Formally, the limiting model exists also for $\lambda = \infty$, since the range of f is compactified.

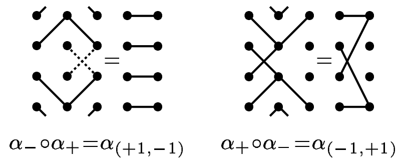
1e1 Example. Define

$$\alpha_-, \alpha_+ : \mathbb{Z} + \frac{1}{2} \to \mathbb{Z} + \frac{1}{2} ,$$
$$\alpha_-(x) = x - 1 , \quad \alpha_+(x) = x + 1 \quad \text{for } x \in \big(\mathbb{Z} + \tfrac{1}{2}\big) \cap (0, \infty) ,$$
$$\alpha_-(-x) = -\alpha_-(x) , \quad \alpha_+(-x) = -\alpha_+(x) .$$



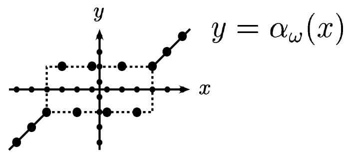
These are not invertible functions; α_- is not injective, α_+ is not surjective; well, we do not need to invert them, we need their compositions:

$$\alpha_\omega = \alpha_{\tau_n(\omega)} \circ \dots \circ \alpha_{\tau_1(\omega)} .$$



All compositions belong to a two-parameter set of functions $h_{a,b}$,

$$\alpha_\omega(x) = h_{a,b}(x) = \begin{cases} x + a & \text{for } x \ge b, \\ x - a & \text{for } x \le -b, \\ (-1)^{b-x}(a+b) & \text{for } -b \le x \le b; \end{cases}$$
$$b, a + b \in \big(\mathbb{Z} + \tfrac{1}{2}\big) \cap (0, \infty) = \{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \dots\} .$$



Indeed, $\alpha_- = h_{-1,1.5}$, $\alpha_+ = h_{1,0.5}$, and $h_{a_2,b_2} \circ h_{a_1,b_1} = h_{a,b}$ where $a = a_1 + a_2$, $b = \max(b_1, b_2 - a_1)$. Thus, $\alpha_\omega = h_{\alpha(\omega),b(\omega)}$, and we define

$$f_1 : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to \mathbb{R}^2 \times \{-1, +1\} ,$$
$$f_1(\omega) = \bigg(\frac{a(\omega)}{\sqrt{n}}, \frac{b(\omega)}{\sqrt{n}}, (-1)^{b(\omega)-0.5} \bigg) .$$

Though, the function is neither bounded nor real-valued; in order to fit into the framework of 1c we take, say, $\arctan\big(a(\omega)/\sqrt{n}\big)$, $\arctan\big(b(\omega)/\sqrt{n}\big)$, and $(-1)^{b(\omega)-0.5}$. The latter is essential if, say, $\frac{1}{\sqrt{n}}\alpha_\omega(0.5)$ is treated as an 'observable'; you see, $\frac{1}{\sqrt{n}}\alpha_\omega(0.5) = (-1)^{b(\omega)-0.5}\frac{1}{\sqrt{n}}(a(\omega) + b(\omega))$. The limiting model exists, and is quite interesting. We'll return to it later. As before, a random process appears by considering the composition over $k < un$.

Here, finite models are not invariant under permutations of their independent random variables (since the maps do not commute), and the limiting model appears not to be invariant under measure preserving transformations of $[0, 1]$.

Independence present in finite models survives in the limit, provided that the limit is described by a two-parameter random process; we'll return to the point later.

1f Trees, not cubes

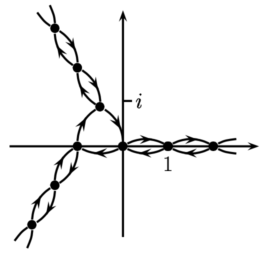
1f1 Example. A particle moves on the sphere S^2. Initially it is at a given point $x_0 \in S^2$. Then it jumps by ε in a random direction. That is, $X_0 = x_0$, while the next random variable X_1 is distributed uniformly on the circle $\{x \in S^2 : |x_0 - x| = \varepsilon\}$. Then it jumps again, to X_2 such that $|X_1 - X_2| = \varepsilon$, and so on. We have a Markov chain (X_k) in discrete time (and continuous space). Let Ω_ε be the corresponding probability space; it may be the space of sequences $(x_0, x_1, x_2, \dots)$ satisfying $|x_k - x_{k+1}| = \varepsilon$, or something else, but in any case $X_k : \Omega_\varepsilon \to S^2$. We choose $\varepsilon_n \to 0$ (say, $\varepsilon_n = 1/n$), take $\Omega_n = \Omega_{\varepsilon_n}$ and define $f_u : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to S^2$ by

$$f_u(\omega) = X_k(\omega) \quad \text{for } \varepsilon_n^2 k \le u < \varepsilon_n^2 (k+1)\,, \quad \omega \in \Omega_n\,.$$

Of course, the limiting model is the Brownian motion on the sphere S^2.

In contrast to previous examples, here Ω_n is not a product; the n-th model does not consist of independent random variables. Though, we can parametrize these Markov transitions by independent random variables; however, there is a lot of freedom in doing so; no one among the parametrizations may be called canonical. The same holds for the limiting model. The Brownian motion on S^2 can be driven by the Brownian motion on R^2 according to some stochastic differential equation, but the latter involves a lot of freedom.

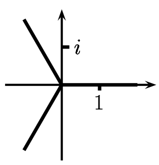
1f2 Example. Consider the random walk on such an oriented graph:



A particle starts at 0 and chooses at random (with probabilities 1/2, 1/2) one of the two outgoing edges; and so on (you see, exactly two edges go out of any vertex). Such $(Z_0, Z_1, \dots)$ is known as the simplest spider walk. It is a complex-valued martingale. The set Ω_n of all n-step trajectories contains 2^n elements and carries its natural structure of a binary tree. (It can be mapped to the binary cube $\{-1, +1\}^n$ in many ways.) We define $f_u : (\Omega_1 \uplus \Omega_2 \uplus \dots) \to \mathbb{C}$ by

$$f_u(\omega) = \frac{1}{\sqrt{n}} Z_k(\omega) \quad \text{for } k \le nu < k+1\,, \quad \omega \in \Omega_n\,.$$

The limiting model is a continuous complex-valued martingale whose values belong to the union of three rays.



The process is known as Walsh's Brownian motion, a special case of so-called spider martingale.

1g Sub-σ-fields

Every example considered till now follows the pattern of 1c; a joint compactification of probability spaces Ω_n, and the limiting Ω. Moreover, Ω_n is usually related to a set T_n (a parameter space, interpreted as time or space), and Ω to a joint compactification T of these T_n.

Example	T_n	T
1b1	$\{1, \frac{1}{2}, \dots, \frac{1}{n}\}$	$\{1, \frac{1}{2}, \frac{1}{3}, \dots\} \cup \{0\}$
1b3	$[\varepsilon_n, 1]$	$[0, 1]$
1d1, 1d2, 1d3, 1e1, 1f1, 1f2	$\{\frac{1}{n}, \frac{2}{n}, \dots, 1\}$	$[0, 1]$

Examples 1b1, 1b3, 1d3 deal (for a finite n) with independent increment processes, taking on their values in a group, namely, 1d3: $\mathbb{R}$ (additive); 1b1: $\{-1, +1\}$ (multiplicative), 1b3: the circle $\{z \in \mathbb{C} : |z| = 1\}$ (multiplicative). Every $t \in T_n$ splits the process in two pieces, the past and the future; in order to keep them independent, we define them via increments, not values.[2] In terms of random signs τ_k (for 1b1, 1d3) it means simply $\{-1, +1\}^n = \{-1, +1\}^k \times \{-1, +1\}^{n-k}$; here k depends on t. The same idea (of independent pieces) is formalized by sub-σ-fields $\mathcal{F}_{0,t}$ (the past) and $\mathcal{F}_{t,1}$ (the future) on our probability space (Ω_m or Ω). Say, for the Brownian motion (1d3), $\mathcal{F}_{0,t}$ is generated by Brownian increments on $[0, t]$, while $\mathcal{F}_{t,1}$ — on $[t, 1]$. Similarly we may define $\mathcal{F}_{s,t}$ for $s < t$, and we have

$$\mathcal{F}_{r,s} \otimes \mathcal{F}_{s,t} = \mathcal{F}_{r,t} \quad \text{whenever } r < s < t\,;$$

it means two things: first, independence,

$$\mathbb{P}\,\big(\,A \cap B\,\big) = \mathbb{P}\,\big(\,A\,\big)\mathbb{P}\,\big(\,B\,\big) \quad \text{whenever } A \in \mathcal{F}_{r,s}, B \in \mathcal{F}_{s,t}\,;$$

and second, $\mathcal{F}_{r,t}$ is generated by $\mathcal{F}_{r,s}$ and $\mathcal{F}_{s,t}$ (that is, $\mathcal{F}_{r,t}$ is the least sub-σ-field containing both $\mathcal{F}_{r,s}$ and $\mathcal{F}_{s,t}$). Such a two-parameter family $(\mathcal{F}_{s,t})$ of sub-σ-fields is called a *factorization* (of the given probability space). Some additional precautions are needed when dealing with semigroups (like 1e1), and also, with discrete time.

Sub-σ-fields $\mathcal{F}_A$ can be defined for some subsets $A \subset T$ more general than intervals, getting

$$\mathcal{F}_A \otimes \mathcal{F}_B = \mathcal{F}_C \quad \text{whenever } A \uplus B = C\,.$$

Models of high symmetry admit arbitrary measurable sets A; models of low symmetry do not. For some examples (such as 1d1, 1d2), a factorization emerges after the limiting procedure.[3]

No factorization at all is given for 1f1, 1f2. Still, the past $\mathcal{F}_{0,t} = \mathcal{F}_t$ is defined naturally. However, the future is not defined, since possible continuations depend on the past. Here we deal with a one-parameter family $(\mathcal{F}_t)$ of sub-σ-fields, satisfying only a monotonicity condition

$$\mathcal{F}_s \subset \mathcal{F}_t \quad \text{whenever } s < t\,;$$

such $(\mathcal{F}_t)$ is called a *filtration*.

[2] In fact, the process of 1b1 has also independent values (not only increments); but that is irrelevant.

[3] For 1d2, some factorization is naturally defined for Ω_n, but is lost in the limiting procedure, and a new factorization emerges.